U.S. Army Public Health Command

Public Health Report

Statistical Process Control Charts for Public Health Monitoring

PHR No. S.0023112

General Medical: 500A, Public Health Data

December 2014

USAPHC

Epidemiology and Disease Surveillance Portfolio
Injury Prevention Program

Prepared by Participants of the Injury Prevention Program:

Dr. Anna Schuh
Dr. Michelle Canham-Chervak

Table of Contents

Statistical Process Control Charts for
Public Health Monitoring
PHR No. S.0023112

1 Summary

1.1 Purpose

To describe details of control charting methods for monitoring injuries among Active Duty Army Soldiers.

1.2 Overview

Statistical process control charts have recently been used for public health monitoring, predominantly in healthcare and hospital applications such as the surveillance of patient wait times or the frequency of surgical failures [e.g., 1-10]. Because the frequency of safety incidents like industrial accidents and motor vehicle crashes will follow a similar probability distribution, the use of control charts for their surveillance has also been recommended [11-15]. These control chart uses can be extended to military applications, such as monitoring active duty Army injuries.

2 References

See Appendix A for a complete list of reference information.

3 Background

3.1 Historical Application

Statistical process control (SPC) monitoring is a process that has been used in many applications to understand past process performance, remove existing sources of natural and unnatural variability, and identify any new sources of variability [1]. Control charts are SPC tools that have traditionally been used in manufacturing to monitor process defects. The variables being monitored can be count data such as the number of non-conforming parts produced in the past month, or numerical quantities such as the length measurement of a part being produced [2].

Basic control charts were first proposed by Walter A. Shewhart in 1929 and are, therefore, known as Shewhart charts. The various Shewhart charts, the appropriate conditions for their use, and some healthcare examples are given in Table 1.

Table 1. Types of Shewhart Charts (adapted from [1])

Type of control chart	Probability distribution	Measurements to apply to	Healthcare examples
X-bar and S chart	normal (Gaussian)	Continuous measurements with a "bell shape"	Length of patient wait times Procedure durations
np-chart	binomial	Total number of dichotomous (yes/no) cases that result in a certain outcome. Only one outcome per sampling unit may be observed (e.g., a surgery either develops an infection or it does not). Assumes a constant sample size/population.	Number of surgeries that develop an infection Number of patients readmitted to the hospital
p-chart	binomial	Fraction of dichotomous (yes/no) cases that result in a certain outcome. Only one outcome per sampling unit may be observed (e.g., a surgery either develops an infection or it does not). The fraction cannot exceed 1.00. Sample size/population may change from sample to sample.	Fraction of surgeries that develop an infection Fraction of patients readmitted to the hospital
c-chart	Poisson	Total number of incidents. There may be more than one event per sampling unit (e.g., more than one fall per patient). Assumes a constant sample size/population.	Number of patient falls Number of ventilator associated pneumonias
u-chart	Poisson	Rate of incidence. There may be more than one event per sampling unit (e.g., more than one fall per patient). The rate may exceed 1.00. Sample size/population may change from time period to time period.	Number of patient falls per 100 patient days Number of ventilator associated pneumonias per 100 ventilator days

3.2 Parameters

Control charts apply upper and lower thresholds called control limits to ensure the stability of a monitored variable. The control limits are typically set at three standard deviations (3σ) above and below the historical Phase I average, because the probability of a data point falling inside these limits if the process remains in statistical control is 99.73% [1]. This can also be thought of as each data point having a false alarm probability of 0.27%. The upper and lower control limits, UCL and LCL, are stated as:

$$UCL = \bar{\mu} + 3\sigma \quad [1]$$

$$LCL = \bar{\mu} - 3\sigma \quad [2]$$

where $\bar{\mu}$ is the historical average or mean.

A baseline dataset of typical data is used to establish the UCL, LCL, historical mean, and historical standard deviation parameters. This baseline is referred to as Phase I. The consideration of as many Phase I historical data points as possible is desired in order to establish the most accurate parameters [1], but at least 20-25 historical data points are typically used to reliably conclude historical process stability [1, 2, 3]. When needed, provisional control limits based on fewer data points can be used and refined over time [4]. The causes of any Phase I points outside the established control limits should be investigated. If the cause is identified and addressed, the data point should be removed so that it does not influence the Phase I parameters. It is often difficult to identify specific assignable causes for out-of-control Phase I data points in health-related data, so obtaining a state of statistical control in the baseline dataset can be a demanding task [16]. The overall false alarm probability for a dataset with 25 historical data points would be fairly low, $1-(.9973)^{25} = 6.5\%$.

New data points are actively monitored as they are observed in Phase II [5, 8]. In Phase II, when a data point surpasses established upper or lower thresholds, the chart will "signal", indicating that the process is beyond its statistical control limits [5]. When the chart is monitoring negative events like injuries, a signal above the upper limit indicates a significant increase in the frequency. A signal below the lower control limit indicates a significant improvement, or decrease in the frequency.

4 Current Application for Monitoring of Military Injuries

4.1 U-Chart

A u-chart was chosen to monitor the rate of military injuries. This chart is appropriate because the Poisson-distributed injury rates are collected and reported in aggregate time periods (e.g., annually, quarterly, monthly), and it is possible that more than one injury will be experienced per person. As shown in Table 1, the only Shewhart chart that accounts for these properties is the u-chart. In a given time period i, the upper and lower control limits can be calculated by

$$UCL_i = \bar{\mu} + 3\sigma_i \quad [3]$$

$$LCL_i = \bar{\mu} - 3\sigma_i \quad [4]$$

where $\bar{\mu}$ is the historical Phase I mean, the standard deviation σ_i can be calculated as

$$\sigma_i = \sqrt{\frac{\bar{\mu}}{n_i}}, \quad [5]$$

and n_i is the total person time for the time period i. Therefore, because the total soldier person-time will change from time period to time period, σ_i, UCL_i, and LCL_i will also change from time period to time period [2]. $\bar{\mu}$ should be calculated using all Phase I data and used for every Phase I and Phase II calculation of σ_i, UCL_i, and LCL_i.

4.2 Overdispersion

A common problem with large datasets such as military injury data is the existence of overdispersion, or greater variation than would usually be statistically expected, which can lead to tight control limits and many false signals [8]. To address the potential for overdispersion, Laney (2002) suggested the introduction of a correction factor to account for between-group standard deviation [17]. The correction factor allows for consideration of both between-group and within-group standard deviations. The new control limits are given as:

$$UCL_i = \bar{\mu} + 3\sigma_{within\ group,i}\sigma_{between\ groups} \quad [6]$$

$$LCL_i = \bar{\mu} - 3\sigma_{within\ group,i}\sigma_{between\ groups} \quad [7]$$

Laney's equation for between-group standard deviation is:

$$\sigma_{between\ groups} = \frac{1}{k-(1.128)}\sum_{i=2}^{k} abs\left(\frac{\mu_i - \bar{\mu}}{\sqrt{\frac{\bar{\mu}}{n_i}}} - \frac{\mu_{i-1} - \bar{\mu}}{\sqrt{\frac{\bar{\mu}}{n_{i-1}}}}\right) \quad [8]$$

where k represents the total number of Phase I observations. This equation for between-group standard deviation essentially calculates the average difference between rates in subsequent periods. Like $\bar{\mu}$, $\sigma_{between\ groups}$ should be calculated using all Phase I data and used in every Phase I and Phase II calculation of UCL_i and LCL_i. Combining Equations 5, 9, and 10 into Equations 6 and 7, the detailed equations for UCL and LCL that correct for overdispersion are:

$$UCL_i = \bar{\mu} + 3\sqrt{\frac{\bar{\mu}}{n_i}}\left[\frac{1}{k-(1.128)}\sum_{i=2}^{k} abs\left(\frac{\mu_i - \bar{\mu}}{\sqrt{\frac{\bar{\mu}}{n_i}}} - \frac{\mu_{i-1} - \bar{\mu}}{\sqrt{\frac{\bar{\mu}}{n_{i-1}}}}\right)\right] \quad [9]$$

$$LCL_i = \bar{\mu} - 3\sqrt{\frac{\bar{\mu}}{n_i}}\left[\frac{1}{k-(1.128)}\sum_{i=2}^{k} abs\left(\frac{\mu_i - \bar{\mu}}{\sqrt{\frac{\bar{\mu}}{n_i}}} - \frac{\mu_{i-1} - \bar{\mu}}{\sqrt{\frac{\bar{\mu}}{n_{i-1}}}}\right)\right] \quad [10]$$

4.3 FORSCOM and TRADOC Examples

Equations 9 and 10 were used to create control charts for active duty Army injuries by installation in the Strategic Management System (SMS). A Phase I u-chart for injury rates among Active Duty personnel at the United States Army Forces Command (FORSCOM) headquarters at Fort Bragg is shown in Figure 1. The 2007-2013 quarterly injury rates were provided by the Armed Forces Health Surveillance Center (AFHSC). The definition of injury is consistent with recommendations for monitoring military injuries [18]. Beginning with 1st quarter 2014, the Phase II data will be monitored using the parameters established in Phase I. We note that there is one out-of-control point in the Phase I data, for the 1st quarter of 2012. This increase in injury rates occurred Army-wide, but the specific cause of this increase could not be identified. Therefore, the outlier will not be removed from Phase I data. Efforts should continue to be made to identify the cause of this increase; the false alarm probability for this dataset is only 7.3%. Lower injury rates in the 4th quarter of most years should also be noted. This may be attributed to the relatively high number of military holidays and associated leave during 4th quarter. Seasonal variations in injury rates have been demonstrated in Army basic training populations [19]. Previous data has suggested that injury rates are higher in summer months (3rd quarter) whereas this data exhibits the highest rates on average during 1st quarter. Control charting provides a comprehensive visualization of these dynamic trends.

Linear regression has been recommended for community health analysis as a method for identifying trends in data [20]. A linear regression of the 2007-2013 Fort Bragg data points does not reveal a statistically significant linear trend.

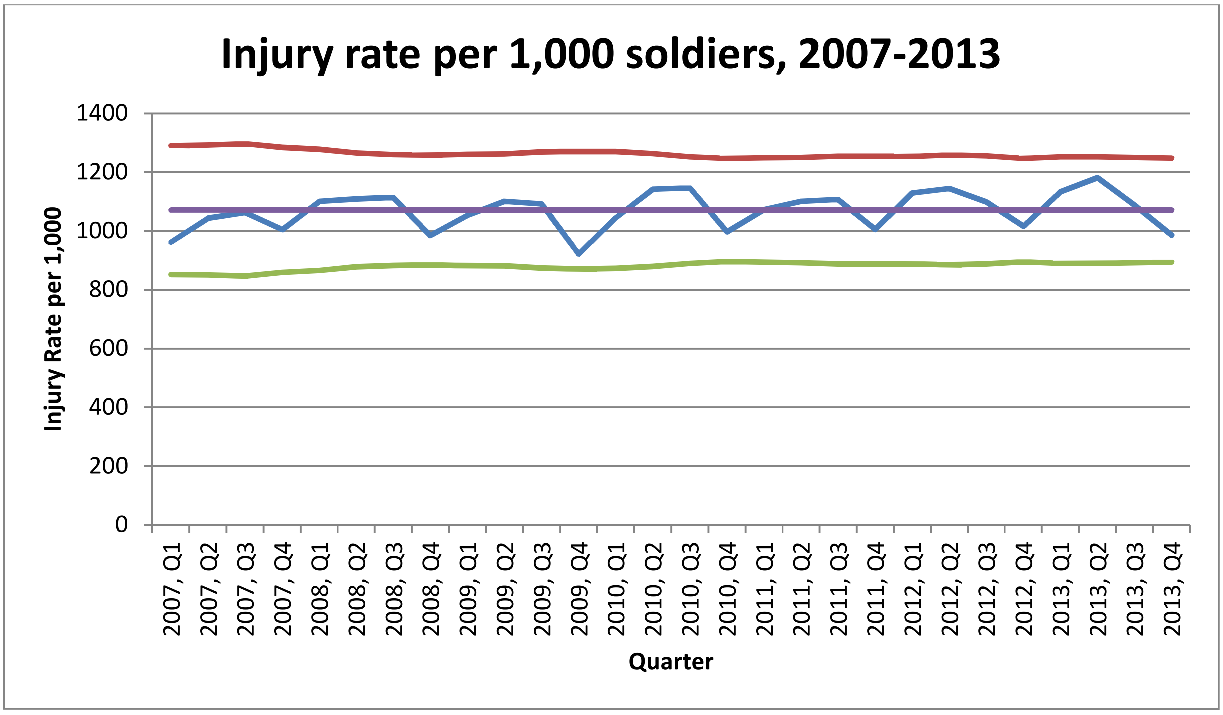

Figure 1. Phase I Control Chart for Quarterly Injury Rates at Fort Bragg, 2007-2013

Control charts for the injury rates at the Army Training and Doctrine Command (TRADOC) basic training sites were also investigated (Fort Jackson, Fort Benning, Fort Sill, and Fort Leonard Wood). These rates are shown in the control chart in Figure 2.

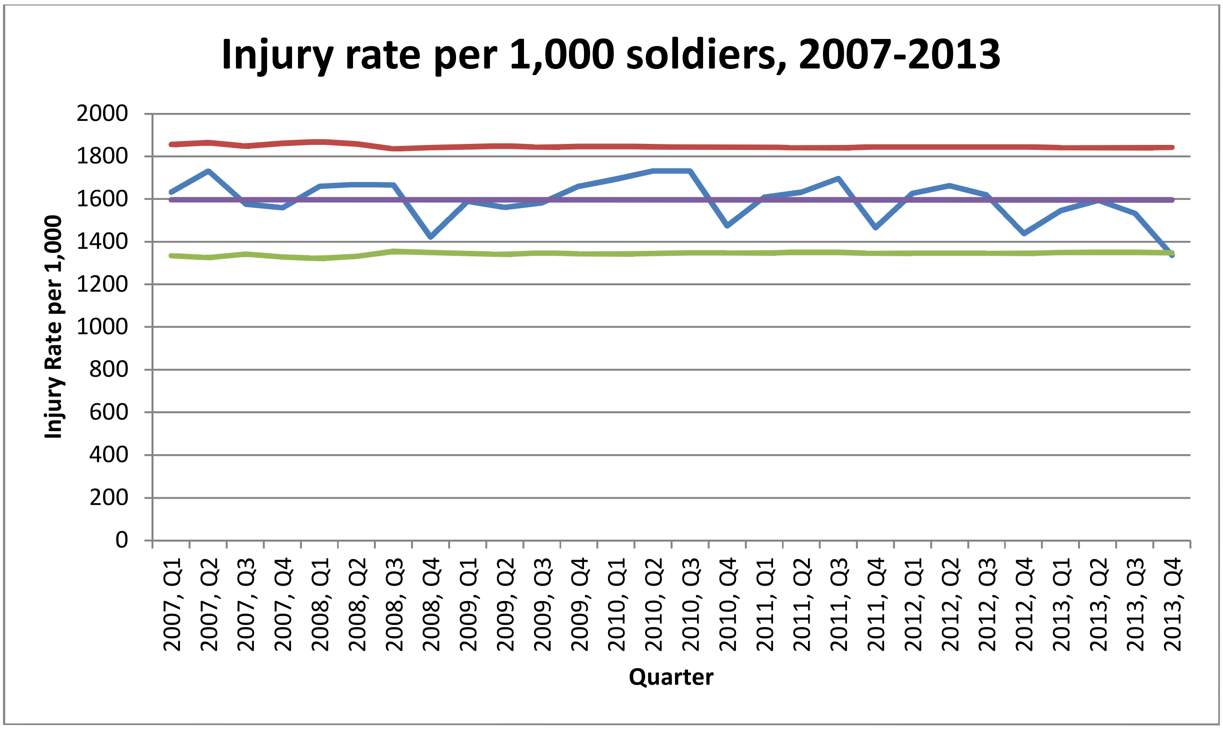

Figure 2. Phase I Control Chart for Quarterly Injury Rates at TRADOC Basic Training Installations, 2007-2013

While the combined injury rates at these four training installations are higher than the rates at the single installation represented in Figure 1, the rates at the training sites experienced a statistically significant linear decrease ($p=0.01$) over the 28-quarter Phase I baseline period. If this tendency continues in Phase II, consistently low rates below the LCL will indicate statistically significant improvements in injury rates. If this is the case, the successful injury prevention strategies used in TRADOC training installations should be identified and implemented in other installations like Fort Bragg as well.

There are many similarities in the two data sets, including a relatively high injury rate in the 1st quarter of 2012 and regularly low injury rates during 4th quarters. However, a paired-samples t-test reveals that the two trends are significantly different ($p<0.001$) since the TRADOC rates are linearly decreasing and the FORSCOM rates are not. T-tests are used compare the regression trends in two datasets and determine whether they are statistically similar [20].

5 Considerations for Future Work

Many extensions to basic Shewhart charts have been recommended for specific applications. These alternate methods often provide more efficient signaling rates, but may also require more rigorous statistical computation. As public health monitoring for military populations progresses, some potential future considerations are listed below.

- If a signal indicates an improvement in the rate of a negative event, the potential causes of this improvement should be identified and reinforced. If improved data continue to be observed, the control limits should be redefined and monitoring should be reset.

- 2σ warning limits could be added to encourage earlier investigation into potential causes of deviations beyond control limits. Since data points will see a higher false alarm rate with these lower limits (5% instead of 0.07%), a point outside of them should not be treated as seriously as a point outside the standard 3σ limits.

- In the 1950s, Western Electric developed decision rules for detecting nonrandom patterns in data, in addition to the signal outside the 3σ control limits, which could also be adapted for monitoring Army injury rates. These additional rules include [2]:

    - Two out of three consecutive points plot beyond 2σ warning limits
    - Four out of five consecutive points plot at a distance of 1σ or beyond from the center line, or
    - Eight consecutive points plot on one side of the center line.

- A lower degree of data aggregation (e.g., monthly data rather than quarterly or annual data) is desired to signal process shifts more quickly and reduce data loss [1, 6]. Furthermore, lower degrees of aggregation would not require historical data that is as dated; for example, only 2 years of monthly data would be needed to provide 24 baseline data points, rather than the 6 years of quarterly data needed to provide 24 data points. In addition, if near-real-time monitoring of time-between-events data can be achieved, this Weibull- or exponentially-distributed data would provide better control chart performance than current methods (aggregated Poisson counts) [21-23].

- Cumulative sum (CUSUM) and exponentially weighted moving average (EWMA) control charts are often used with Phase II data. These charts have been shown to more quickly detect small changes than traditional Shewhart charts. There have been several applications of CUSUM charts in public health applications like monitoring surgical failures or motor vehicle accidents [e.g., 4, 7, 11, 14-15].

- It has been suggested that overdispersed count data should be modeled as a negative binomial process instead of a Poisson process [10, 24]. Since there is no Shewhart chart for this distribution, a CUSUM or EWMA chart would be required.

- Risk adjustment for health data has been applied when monitoring variables that can be conditional on an individual's personal degree of risk [e.g., 5, 7, 9-10]. For example, Steiner "risk-adjusted" surgical fatality data based on a variety of factors, so that fatalities in surgical patients with a lower pre-determined mortality risk were considered more significant and were weighted more than fatalities in patients who entered surgery with a higher fatality risk [7]. In this way, risk adjustment allows for some degree of standardization and consideration of situational influence.

- There are some practical challenges to consider when implementing statistical process control tools in health-related fields. Some concerns that may arise when monitoring health data with control charts include, but are not limited to, the availability of appropriate data; seasonality or other autocorrelation; uncertainty in control chart selection; and incorrect interpretation of chart data [9, 10, 17]. These challenges should be considered as Army injury control charts are refined.

6 Point of Contact

The USAPHC IPP is the point of contact for this project, e-mail usarmy.apg.medcom-phc.mbx.injuryprevention@mail.mil, or phone number 410-436-4655, DSN 584-4655. Specific questions may be directed to author(s) listed at the front of this report.

Approved:

DR. BRUCE JONES
Program Manager
Injury Prevention Program

Appendix A

References

1. Benneyan, J.C. 2008. The Design, Selection, and Performance of Statistical Control Charts for Healthcare Process Improvement. *International Journal of Six Sigma and Competitive Advantage* 4(3): 209-39.
2. Montgomery, D. C. 2009. Introduction to Statistical Quality Control. New York: Wiley.
3. Benneyan, J.C., Lloyd, R.C., and Plsek, P.E. 2003. Statistical Process Control as a Tool for Research and Healthcare Improvement. *Quality and Safety in Health Care* 12: 458-64.
4. Mohammed, M. A., Worthington, P., and Woodall, W.H. 2008. Plotting Basic Control Charts: Tutorial Notes for Healthcare Practitioners. *Quality and Safety in Health Care* 17(2): 137-45.
5. Grigg, O.A., Farewell, V.T., and Spiegelhalter, D.J. 2003. Use of Risk-Adjusted CUSUM and RSPRT Charts for Monitoring in Medical Contexts. *Statistical Methods in Medical Research* 12: 147-70.
6. Sonesson, C. and Bock, D. 2003. A Review and Discussion of Prospective Statistical Surveillance in Public Health. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 166(1): 5-21.
7. Steiner, S. H. 2000. Monitoring Surgical Performance Using Risk-adjusted Cumulative Sum Charts. *Biostatistics* 1(4): 441-52.
8. Woodall, W.H. 2006. The Use of Control Charts in Health-Care and Public-Health Surveillance. *Journal of Quality Technology* 38(2): 89-134.
9. Woodall, W.H, Grigg, O.A., and Burkom, H.S. 2010. Research Issues and Ideas on Health-Related Surveillance. *Frontiers in Statistical Quality Control* 9: 145-55.
10. Woodall, W.H., Adams, B.M., and Benneyan, J.C. 2011. The Use of Control Charts in Healthcare. Statistical Methods in Healthcare, F. Faltin, R. Kenett, F. Ruggeri, eds., Wiley.
11. Adekeye, K.S. and Aluko, O.S. 2012. Design of Cusum Scheme for Monitoring Road Accident Fatalities. *Open Journal of Statistics* 2(2): 213-18.
12. Cournoyer, M. E., Garcia, V.E., Gallegos, U.F., and Wilburn, D.W. 2011. Investigation of Injury/Illness Data at a Nuclear Facility. *Journal of Chemical Health and Safety* 18(5): 17-25.
13. Guria, J. and Mara, K. 2000. Monitoring Performance of Road Safety Programmes in New Zealand. *Accident Analysis & Prevention* 32(5): 695-702.
14. Lucas, J. M. 1985. Counted Data CUSUM's. *Technometrics* 27(2): 129-44.
15. Wu, Z., Liu, Y., He, Z. and Khoo, M.B.C. 2010. A Cumulative Sum Scheme for Monitoring Frequency and Size of an Event. *Quality and Reliability Engineering International* 26(6): 541-54.
16. Jordan, V. and Benneyan, J.C. 2011. Common Challenges and Pitfalls Using SPC in Healthcare. Statistical Methods in Healthcare, F. Faltin, R. Kenett, F. Ruggeri, eds., Wiley.
17. Laney, D. B. 2002. Improved Control Charts for Attributes. *Quality Engineering* 14(4): 531-37.
18. DoD Military Injury Metrics Working Group White Paper. 2002.
19. Knapik, J.J., Canham-Chervak, M., Hauret, K., Laurin, M. J., Hoedebecke, E., Craig, S., and Montain, S.J. 2002. Seasonal Variations in Injury Rates During US Army Basic Combat Training. *Annals of Occupational Hygiene* 46(1): 15-23.
20. Remy, L.L., Clay, T., and Oliva, G. 2005. Do We Have a Linear Trend? A Beginner's Approach to Analysis of Trends in Community Health Indicators. *Family Health Outcomes Project Report*.

21. Schuh, A., Woodall, W.H., and Camelio, J.A. 2013. The Effect of Aggregating Data When Monitoring a Poisson Process. *Journal of Quality Technology* 45(3): 260-72.
22. Schuh, A., Camelio, J.A., and Woodall, W.H. 2014. Control Charts for Accident Frequency: A Motivation for Real Time Accident Surveillance. *International Journal for Injury Control and Safety Promotion* 21(2): 154-62.
23. Shafae, M.S., Dickinson, R.M., Woodall, W.H., and Camelio, J.A. 2014. Cumulative Sum Control Charts for Monitoring Weibull-distributed Time Between Events. *Quality and Reliability Engineering International*, doi 10.1002/qre.1643.
24. Szarka, J. L., III, Gan, L., and Woodall, W. H. 2011. Comparison of the Early Aberration Reporting System (EARS) W2 Methods to an Adaptive Threshold Method. *Statistics in Medicine* 30(5): 489-504.